REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	5/29/98	Final- 1 April 97- 31 May 98

4. TITLE AND SUBTITLE

Number Theoretic Methods in Parameter Estimation

5. FUNDING NUMBERS

(g)

ONR Grant Number

N0014-97-1-0566

6. AUTHOR(S)

Stephen D. Casey

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

American University
Department of Mathematics and Statistics
4400 Massachusetts Ave., N.W.
Washington, D.C. 20016-8050

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

19980617 104

11. SUPPLEMENTARY NOTES

See appendix for publication and awards

a. DISTRIBUTION / AVAILABILITY STATEMENT

unlimited

12. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We have combined results from analysis and number theory with statistical signal processing to develop new results in parameter estimation. In particular, our developments in the theory on the Riemann Zeta Function and algorithms on extensions of Euclidean domains have led to new computationally straightforward algorithms for parameter estimation from sparse, noisy data. Robust versions have been developed that are stable despite significant jitter noise and the presence of arbitrary outliers. We have continued the development of the theory, including the development computationally straightforward techniques for spectral analysis of a very broad class of periodic processes, including procedures so that estimates achieve the Cramer-Rao bound. We have extended these techniques to the complete analysis of zero-crossing data and multiply periodic point processes, including the recovery of the fundamental period(s), phase information, the multiples of the periods, and the deinterleaving of the data. The algorithms will work on data from currently deployed sonar, radar, and communication systems. We have also applied our techniques to other data sets containing sparse noisy information generated by a periodic process, e.g., the geometric pattern generated by minefield placement. We also briefly report on our work on multichannel deconvolution and our work on derivatives.

14. SUBJECT TERMS

radar, sonar, parameter estimation, modified Euclidean algorithm periodogram, Riemann Zeta Function, least squares.

15. NUMBER OF PAGES

18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	UL

Number Theoretic Methods in Parameter Estimation: Theory and Application
ONR Grant Number N00014–97–1–0566
FINAL REPORT

Stephen D. Casey
Department of Mathematics and Statistics
American University
4400 Massachusetts Ave., N.W.
Washington, DC 20016–8050
scasey@american.edu

May 29, 1998

Abstract

Problems in harmonic analysis and synthesis are intertwined with their applications in signal and image processing. Some recent advances in the theory of this analysis have used number theory to extend existing theories (e.g., sampling theory, fast computations) and develop new approaches to problems (e.g., interpolation). An area that has benefited from the blend of number theory and harmonic analysis is that of parameter estimation. This is the main area of focus for our work in this grant.

Our developments in the theory on the Riemann Zeta Function and algorithms on extensions of Euclidean domains have led to new computationally straightforward algorithms for parameter estimation from sparse, noisy data. Robust versions have been developed that are stable despite significant jitter noise and the presence of arbitrary outliers.

We have continued the development of the theory, including the development computationally straightforward techniques for spectral analysis of a very broad class of periodic processes, including procedures so that estimates achieve the Cramer-Rao bound. We have extended these techniques to the complete analysis of zero-crossing data and multiply periodic point processes, including the recovery of the fundamental period(s), phase information, the multiples of the periods, and the deinterleaving of the data. The algorithms will work on data from currently deployed sonar, radar, and communication systems. We have also applied our techniques to other data sets containing sparse noisy information generated by a periodic process, e.g., the geometric pattern generated by minefield placement.

The mathematical techniques used in our work on parameter estimation was also applicable to another fundamental area of signal and image processing, that of deconvolving a given signal from the sensor that gathers it. This approach has been labeled *multichannel deconvolution.* It circumvents the ill-posedness in the single channel approach by overdetermining the input signal using a system of *strongly coprime* impulse responses. We have also continued the development of several aspects of this theory, including applying our theory to develop systems in which complete signal information can be recovered. We also have extended the theory to more general system models, and create a basis in which to develop the multichannel theory which is more amenable to discretization. We also have coordinated the theory with filtering systems for the removal of noise and/or other unwanted information, and with irregular sampling theory and wavelet and Gabor analysis. Our work on multichannel deconvolution has also led to a new approach to sampling theory. We have developed specific examples of non-commensurate sampling lattices, and used a generalization of B. Ya. Levin's sine-type functions to develop interpolating formulae on these lattices. We are exploring an extension of these ideas to create non-commensurate systems of wavelet bases.

We close with an additional general mathematical result. As is well known, a function f with a positive derivative Df on an interval is increasing on that interval. We explored the extent to which the hypotheses of the Mean Value Theorem can be weakened and f still shown to be increasing. By constructing counterexamples using Cantor sets, we showed that the assumption $Df > 0$ *a.e.* does not imply that f is increasing. Then we showed that if $Df > 0$ except on a countable subset of an interval, f is increasing. We call this the Countable Exceptional Set Theorem. This theorem is generalized by the Goldowsky-Tonelli Theorem, which tells us that if Df exists except on a countable subset of an interval and $Df > 0$ *a.e.*, then f is increasing. However, we then showed, in a very natural sense, that Goldowsky-Tonelli is a vacuous extension of the Countable Exceptional Set Theorem.

1 Introduction: An Overview of Results

Problems in harmonic analysis and synthesis are intertwined with their applications in signal and image processing. Some recent advances in the theory of this analysis have used number theory to extend existing theories (e.g., sampling theory, fast computations) and develop new approaches to problems (e.g., interpolation). An area that has benefited from the blend of number theory and harmonic analysis is that of parameter estimation. As discussed in our proposal, our work for this grant focused on parameter estimation. However, we also did some work in additional areas and include a brief discussions of these results.

Our developments in the theory on the Riemann Zeta Function and algorithms on extensions of Euclidean domains have led to new computationally straightforward algorithms for parameter estimation from sparse, noisy data. Robust versions have been developed that are stable despite significant jitter noise and the presence of arbitrary outliers.

We[1] have continued the development of the theory. Particular items include:

- **i.)** Developed computationally straightforward techniques for spectral analysis of a very broad class of periodic processes, including procedures so that estimates achieve the Cramer-Rao bound [21, 22, 23, 24, 41, 43]. Our techniques of parameter estimation fit models of currently deployed sonar, radar, and communication systems.
- **ii.)** Extended these techniques to the complete analysis of multiply periodic point processes, including the recovery of the fundamental period(s), phase information, the multiples of the periods, and the deinterleaving of the data [21, 24].
- **iii.)** Employed these techniques to estimating the period of a sinusoid from zero-crossing data. These techniques are applicable to sparse data sets on which conventional techniques break down. This work includes two new derivations (analytic and geometric) of the Tretter approximation [42, 44].
- **iv.)** Apply our techniques to other data sets containing sparse noisy information generated by a periodic process, e.g., the geometric pattern generated by minefield placement [36].

Parameter estimation problems in periodic point processes arises in radar pulse repetition interval (PRI) analysis, in bit synchronization in communications, in neurology and astronomy, and many other scenarios. Again, simulation work indicates that our algorithms can be easily implemented into models of actual systems to perform this analysis. This work is directly applicable to Navy technologies. For example, it will provides new procedures for period estimation and/or the deinterleaving of radar and sonar data, and synchronization for frequency hopping communications. It is also applicable to a variety of other problems, from the analysis of general time series data to the detection of geometric regularities, as might arise in minefield detection.

The mathematical techniques used in our work on parameter estimation was also applicable to another fundamental area of signal and image processing, that of deconvolving a given signal from the sensor that gathers it. This approach has been labeled *multichannel deconvolution* [10]–[19]. It circumvents the ill-posedness in the single channel approach by overdetermining the input signal using a system of *strongly coprime* impulse responses $\{\mu_i\}$. We then filter the output of each channel by the deconvolvers $\{\nu_i\}$ which satisfy $\delta = \mu_1 * \nu_1 + \ldots + \mu_n * \nu_n$. This in turn allows for the recovery of the input signal.

We have continued the development of several aspects of this theory, including applying our theory to develop systems in which complete signal information can be recovered. We also have extended the theory to more general system models, and create a basis in which to develop the multichannel theory which is more amenable to discretization. We also have coordinated the theory with filtering systems for the removal of noise and/or other unwanted information, and with irregular sampling theory and wavelet and Gabor analysis. This work is applicable to Navy technologies in that it provides new theoretical basis for the development of active and passive remote sensing systems which gather all possible information in a noisy environment. Moreover, the systems can be designed to do a simultaneous signal recovery and analysis.

Our work on multichannel deconvolution has also led to a new approach to sampling theory. Solutions to the analytic Bezout equation associated with certain multichannel deconvolution problems are interpolation

[1] The editorial we is used throughout the report. The results are by the PI, and where noted, various co-authors.

problems on unions of non-commensurate lattices. These solutions provide insight into how we developed general sampling schemes on properly chosen non-commensurate lattices. We have developed specific examples of non-commensurate sampling lattices, and used a generalization of B. Ya. Levin's sine-type functions to develop interpolating formulae on these lattices. We are exploring an extension of these ideas to create non-commensurate systems of wavelet bases.

We close with an additional general mathematical result. As is well known, a function f with a positive derivative Df on an interval is increasing on that interval. We explored the extent to which the hypotheses of the Mean Value Theorem can be weakened and f still shown to be increasing. By constructing counterexamples using Cantor sets, we [20] showed that the assumption $Df > 0$ *a.e.* does not imply that f is increasing. Then we showed that if $Df > 0$ except on a countable subset of an interval, f is increasing. We call this the Countable Exceptional Set Theorem. This theorem is generalized by the Goldowsky-Tonelli Theorem, which tells us that if Df exists except on a countable subset of an interval and $Df > 0$ *a.e.*, then f is increasing. However, we then showed, in a very natural sense, that Goldowsky-Tonelli is a vacuous extension of the Countable Exceptional Set Theorem.

The report is organized as follows. In section 2, we give a discussion of our work on parameter estimation. Section 3 is a brief discussion of the theory and applications of the deconvolution theory. The section 4 gives an overview of our work with derivatives.

2 Parameter Estimation

2.1 Recent Results

In this section, we gather our results in parameter estimation, which have appeared in several very recent papers. Many of the PI's results on the subject appear in [23] – "Modifications of the Euclidean algorithm for isolating periodicities from a sparse set of noisy measurements," *IEEE Transactions on Signal Processing*, **44** (8), 2260–2272 (1996) – [24] – "Number theoretic methods in parameter estimation," *Proceedings of IEEE Workshop on Statistical Signal and Array Processing*, 406–409 (1996) – [21] – "Sampling issues in least squares, Fourier analytic, and number theoretic methods in parameter estimation," *31st Annual Asilomar Conference on Signals, Systems, and Computers*, 453–457 (1998) – [43] – "On periodic pulse interval analysis with outliers and missing observations," to appear in *IEEE Transactions on Signal Processing*, 31 pp. – and [44] – "Frequency estimation via sparse zero crossings," submitted to *IEEE Transactions on Signal Processing*, 11 pp. Variations on these results can be found in [22], [41], and [43]. Also see [36] – "Detecting regularity in minefields using collinearity and a modified Euclidean algorithm, *Proc. SPIE*, **3079**, 8 pp. (1997).

Modifications of the Euclidean algorithm were presented for determining the period from a sparse set of noisy measurements in [22], [23]. The elements of the set were assumed to be the noisy occurrence times of a periodic event with (perhaps very many) missing measurements. This problem arises in radar pulse repetition interval (PRI) analysis, in bit synchronization in communications, and other scenarios. The algorithms are computationally straightforward and converge quickly. A robust version is developed that is stable despite the presence of arbitrary outliers. We model the set of measurements of a periodic process as follows. We assume our data is a finite set of real numbers

$$S = \{s_j\}_{j=1}^{n}\,, \text{ with } s_j = k_j\tau + \phi + \eta_j, \tag{1}$$

where τ (the period) is a fixed positive real number, the k_j's are non-repeating positive integers, ϕ (the phase) is a real random variable uniformly distributed over the interval $[0, \tau)$, and the η_j's are zero-mean independent identically distributed (iid) error terms. We assume that the η_j's have a symmetric probability density function (pdf), and that $|\eta_j| < \frac{\tau}{2}$ for all j. We develop an algorithm for isolating the period of the process from this set, which we shall assume is (perhaps very) sparse and noisy. In the noise-free case our basic algorithm, given below, is equivalent to the Euclidean algorithm and converges with very high probability given only $n = 10$ data samples, independent of the number of missing measurements. We assume that the original data set is in descending order, i.e., $s_j \geq s_{j+1}$.

Modified Euclidean Algorithm

1.) After the first iteration, append zero.
2.) Form the new set with elements $s_j - s_{j+1}$.
3.) Sort in descending order.
4.) Eliminate elements in $[0, \eta_0]$ from end of the set.
5.) Algorithm is done if left with a single element. Declare $\widehat{\tau} = s_1$. If not done, go to (1.).

Simulation examples demonstrate successful estimation of τ for $n = 10$ with 99.99% of the possible measurements missing. In fact, with only 10 data samples, it is possible to have the percentage of missing measurements arbitrarily close to 100%. There is, of course, a cost, in that the number of iterations the algorithm needs to converge increases with the percentage of missing measurements.

In the presence of noise and false data (or outliers), there is a tradeoff between the number of data samples, the amount of noise, and the percentage of outliers. The algorithm will perform well given low noise for $n = 10$, but will degrade as noise is increased. However, given more data, it is possible to reduce noise effects and speed up convergence by binning the data, and averaging across bins. Binning can be effectively implemented by using an adaptive threshhold with a gradient operator, allowing convergence in a single iteration in many cases. Simulation results show, for example, good estimation of the period from one hundred data samples with fifty percent of the measurements missing and twenty five percent of the data samples being arbitrary outliers.

Our algorithm is based on several theoretical results, which we now present. The first allows for a modification of the basic Euclidean algorithm.

Proposition 2.1 ([23])

$$\gcd(\tau \cdot k_1, \ldots, \tau \cdot k_n) = \tau \cdot \gcd((k_1 - k_2), (k_2 - k_3), \ldots, (k_{n-1} - k_n), k_n) . \tag{2}$$

We then show that our procedure almost surely converges to the period by proving the following very interesting result. Recall that Riemann's Zeta Function is defined on the complex half space $\{z \in \mathbf{C} : \Re(z) > 1\}$ by $\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$. Euler demonstrated the connection of ζ with number theory by showing that

$$\zeta(z) = \prod_{j=1}^{\infty} \frac{1}{1 - (p_j)^{-z}} , \; \Re(z) > 1 ,$$

where $\mathbf{P} = \{p_1, p_2, p_3, \ldots\} = \{2, 3, 5, \ldots\}$ is the set of all prime numbers. In the following, we let $P\{\cdot\}$ denote probability, card$\{\cdot\}$ denote the cardinality of the set $\{\cdot\}$, and let $\{1, \ldots, \ell\}^n$ denote the sublattice of positive integers in $\mathbf{R}^n$ with coordinates c such that $1 \leq c \leq \ell$. Therefore, $N_n(\ell) = \text{card}\{(k_1, \ldots, k_n) \in \{1, \ldots, \ell\}^n : \gcd(k_1, \ldots, k_n) = 1\}$ is the number of relatively prime elements in $\{1, \ldots, \ell\}^n$.

Theorem 2.1 ([23]) *For $n \geq 2$, we have that*

$$\lim_{\ell \to \infty} \frac{N_n(\ell)}{\ell^n} = [\zeta(n)]^{-1} . \tag{3}$$

Therefore, given n ($n \geq 2$) randomly chosen positive integers $\{k_1, \ldots, k_n\}$,

$$P\{\gcd(k_1, \ldots, k_n) = 1\} = [\zeta(n)]^{-1} . \tag{4}$$

The convergence of the procedure to τ is rather quick.

Proposition 2.2 ([23]) *Let $\omega \in (1, \infty)$. Then*

$$\lim_{\omega \to \infty} [\zeta(\omega)]^{-1} = 1 , \tag{5}$$

converging to 1 from below faster than $(1 - 2^{1-\omega})$.

In order to illustrate the behavior of the algorithm consider the following example. Let the set S of equation (1) be generated as follows. Let $\tau = 1$, $n = 100$ data samples, the jumps in the k_j's be randomly selected from a discrete uniform distribution on the interval $[1, 10]$, and the noise be iid and uniformly distributed as $f_\eta(\eta) \sim \mathcal{U}[-0.1, 0.1]$. A data set S was generated according to these parameters and used as input to our algorithm. Consider the results after one iteration, in which the data has been differenced and sorted into descending order, as plotted in Figure 1. The data are clustered into "steps" around integer multiples of $\tau = 1$. That the steps are all of the same approximate length is due to the uniform distribution in the jumps of the k_j's in the original data set S. Other distributions will result in different proportions. From the structure of the data and the assumptions on the noise we know that the data has a mean that is an integer multiple of τ given by $(k_j - k_{j+1})\tau$, with noise symmetrically distributed around, this mean. This suggests isolating each step and averaging the data within each step to reduce noise effects.

A straightforward method for clustering the data is to employ a gradient operator to determine when a step has occurred. After the first iteration (as in Figure 1) the gradient is estimated, with large gradient values indicating a step or "edge" in the data. We have employed a simple estimator by convolving with an impulse response given by $[-1, 0, 1]$. This operator is well known in signal and image processing, e.g., see Jain [29]. A data-adaptive gradient threshold g_0 is selected as 10% of the maximum gradient value, and data points above this threshold are assumed to correspond to the step edges. After the steps have been isolated the step heights are easily found, and the minimum step height, call it $\tilde{\tau}$, is taken as a coarse estimate of τ. Referring to Figure 1, all of the step heights are approximately equal to τ, again due to the original distribution of the jumps in the k_j's used in generating S. We then use $\tilde{\tau}$ to set two thresholds. The first is $\eta_0 = 0.35\tilde{\tau}$, used to define the neighborhood of zero in which data will be eliminated during each iteration. The second we take to be $y_0 = 0.6\tilde{\tau}$, used to segment the steps at each iteration. The segmentation proceeds by searching for jumps in height greater than y_0, and averaging over each segment. The choices of 0.35 and 0.6 are based on extensive simulation experience, and have been more rigorously justified in specific cases. However, performance is reasonably robust to changes in these weights under the various scenarios considered. The averaging produces significant data reduction, and therefore greatly increases the speed of convergence. The gradient operator is applied only as part of the first iteration, the data reduces rapidly with each iteration and precludes use of the gradient operator except as part of the first iteration.

We summarize the foregoing in the following algorithm statement. Again we assume the data is initially sorted in descending order. Recall that appending zero in the first step appends the previous minimum.

Modified Euclidean Algorithm (With Averaging)

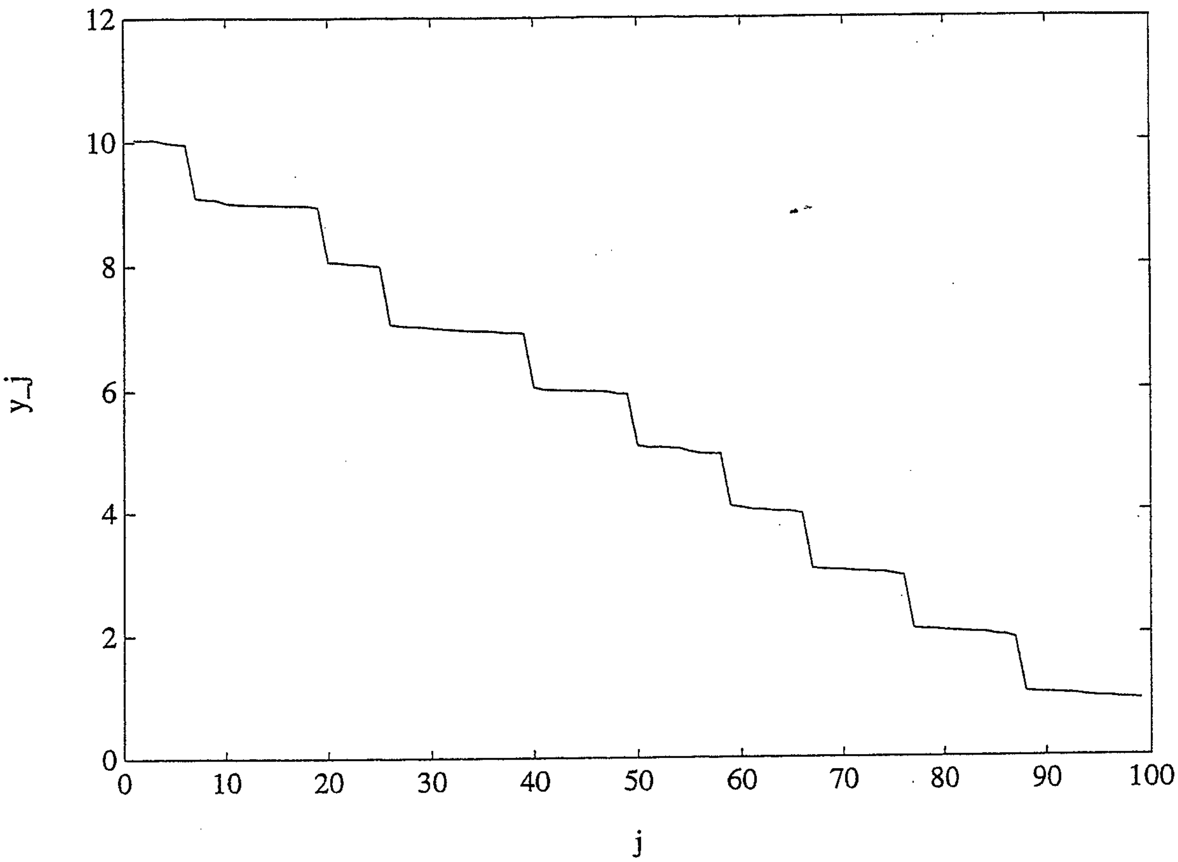
1.) After the first iteration, append zero.
2.) Form the new set with elements $s_j - s_{j+1}$.
3.) Sort in descending order.
4.) On the first iteration, apply gradient and obtain $\tilde{\tau}$, yielding $\eta_0 \cong 0.35\tilde{\tau}$ and $y_0 \cong 0.6\tilde{\tau}$.
5.) Average the data over each step, with steps determined by jumps of height y_0.
6.) Eliminate elements in $[0, \eta_0]$ from end of the set.
7.) Algorithm is done if left with a single element. Declare $\widehat{\tau} = s_1$. If not done, go to **(1.)**.

Further modifications can be found in [23]. The algorithms were tested extensively. Simulation results appear in [22] and [23]. The algorithms perform quite well.

The parameter estimation techniques described above lead to an effective method for periodic pulse interval analysis (see [41], [42]). We assume time is highly resolved and ignore any time quantization error. We are primarily concerned with a single periodic pulse train with (perhaps very many) missing observations that may be contaminated with outliers. Our data model for this case, in terms of the arrival times t_j, is given by (1), with the additional assumption that η_j is zero-mean additive white Gaussian noise. Outliers are included in the data as phase-shifted multiples of another period.

The problem, again, is to recover the period τ and possibly the phase ϕ. The minimum variance unbiased estimates for this linear regression problem take a least-squares form. However, this requires knowledge of the k_j's. We therefore develop a multi-step procedure that proceeds by (i) estimating τ directly, (ii) estimating the k_j's, and (iii) refining the estimate of τ using the estimated k_j's in the least-squares solution. This

Figure 1



1. Plot of example data set after one iteration of the modified Euclidean algorithm of Section 2. The data is sorted in descending order into steps centered around multiples of τ ($\tau = 1$ in this example). The stepwidths are a function of the distribution of the k_j's in the original data set S. The lowest (rightmost) step is centered around the true value of $\tau = 1$.

estimate is shown to perform well, achieving the Cramer-Rao bound in many cases, despite many missing observations and contaminated data. The direct estimate of τ (step (i)) is obtained using the modified Euclidean algorithms described above. While not maximum-likelihood, the modified Euclidean algorithm performs well under difficult conditions.

We now give the maximum likelihood solution and Cramer-Rao bounds for estimating τ and ϕ. Our analysis has led us to work with the data set $\{t_{j+1} - t_j\}_{j=1}^{n-1}$, so as to avoid estimating ϕ (which can be unreliable). Given the sample data set S from (1) we may write

$$\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} = \begin{bmatrix} 1 & k_1 \\ 1 & k_2 \\ \vdots & \vdots \\ 1 & k_n \end{bmatrix} \begin{bmatrix} \phi \\ \tau \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}, \tag{6}$$

where $k_{j+1} > k_j$. In compact form this is

$$\mathbf{t} = X\boldsymbol{\beta} + \boldsymbol{\eta}, \tag{7}$$

where $\boldsymbol{\beta} = [\phi, \tau]^T$ and the definitions of $\mathbf{t}$, $\boldsymbol{\eta}$, and X follow from (6). We eliminate ϕ by forming the differences $y_j = t_{j+1} - t_j = (k_{j+1} - k_j)\tau + (\eta_{j+1} - \eta_j)$, yielding

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix} = \begin{bmatrix} k_2 - k_1 \\ k_3 - k_2 \\ \vdots \\ k_n - k_{n-1} \end{bmatrix} \tau + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{n-1} \end{bmatrix}, \tag{8}$$

where $\delta_j = \eta_{j+1} - \eta_j$. Similar to (7) we may write (8) compactly as

$$\mathbf{y} = X_d\tau + \boldsymbol{\delta}. \tag{9}$$

Equations (7) and (9) are linear regression problems whose least squares solutions yield the minimum-variance unbiased estimate when the noise is zero-mean Gaussian, e.g., see Kay [31]. Generally, use of (9) is preferred for estimating τ, avoiding estimation of ϕ which has high variance. The solution to (9) corresponds to maximum-likelihood estimation and takes the form of an least squares estimate

$$\widehat{\tau} = (X_d^\tau R_\delta^{-1} X_d)^{-1} X_d^\tau R_\delta^{-1}\mathbf{y}, \tag{10}$$

where $R_\delta = E[\boldsymbol{\delta}\boldsymbol{\delta}^\tau]$. We have assumed white noise so $R_\delta = \sigma_\eta^2 \tilde{R}_\delta$ where $\tilde{R}_\delta$ has 2's on the main diagonal, -1's on the first upper and lower diagonals, and zeros elsewhere. In general R_δ is full rank and its inverse can be expressed element-wise as $[R_\delta^{-1}]_{ij} = \min(i,j) - ij/n$, and is therefore easily computed. Although optimal, use of (10) requires knowledge of X_d. This is not a problem if there are no missing observations for then $k_j = j$ for $j = 1, 2, \cdots n$. However, when observations are arbitrarily missing then the k_j's are not known in general, and one is faced with more unknowns than equations in (9).

The pdf of the observations $\mathbf{t}$ in (1) is multivariate Gaussian, leading to the Cramer-Rao bound for (10)

$$\operatorname{var}\{\tau - \widehat{\tau}\} \geq \sigma_\delta^2 (X_d^\tau \tilde{R}_\delta^{-1} X_d)^{-1}, \tag{11}$$

with $\sigma_\delta = 2\sigma_\eta$. Generally, the Cramer-Rao bound is reduced for smaller σ_η^2. Also, for fixed n, it is reduced when the spread of the k_j's increases.

Now, if τ were known then X_d could be estimated using $(1/\tau)\,\mathbf{y}$. Ideally, this estimate is composed of positive integers, but imperfect knowledge of τ and the presence of noise will generally yield an estimate of X_d that has non-integer components. We therefore estimate X_d via

$$\widehat{X}_d = \operatorname{round}\left[\frac{1}{\widehat{\tau}_{MEA}}\,\mathbf{y}\right], \tag{12}$$

where $\widehat{\tau}_{MEA}$ is the estimate of τ obtained via the modified Euclidean algorithm, and round$[\cdot] = \lfloor \cdot + \frac{1}{2} \rfloor$ is rounding to the nearest integer. A refined estimate of τ is then obtained by using $\widehat{X}_d$ in (10) yielding

$$\widehat{\tau} = (\widehat{X}_d^T R_\delta^{-1} \widehat{X}_d)^{-1} \widehat{X}_d^T R_\delta^{-1} \mathbf{y}. \tag{13}$$

This result approaches the optimal minimum variance performance when $\widehat{X}_d$ is close to X_d. The refinement algorithm is summarized as follows.

Refined Estimation Algorithm

1.) Estimate τ via the modified Euclidean algorithm, calling this estimate $\widehat{\tau}_{MEA}$.
2.) Estimate $\widehat{X}_d$ via (12).
3.) Refine the estimate of τ using $\widehat{X}_d$ in (13), calling this estimate $\widehat{\tau}$.

Performance analysis of the estimate $\widehat{\tau}_{MEA}$ depends not only on the distribution of the noise η_j, but also on the distribution of the k_j's. We have completed this analysis for some specific cases in [41]. We also compare the estimates to Cramer-Rao bounds via Monte Carlo simulation, revealing the very good performance of the algorithm with many missing observations and contaminated data. These results can be found in [41], [42].

We can also apply our parameter estimation procedures to frequency estimation via sparse zero crossings. Estimation of the frequency of a single complex sinusoid in Gaussian noise is a fundamental problem in signal processing. We have addressed the problem [42, 44], using only very sparse noisy zero-crossings with the presence of outliers. The techniques presented in these papers require access to zero-crossing times only. The approach relies on the Tretter approximation, which is valid for high SNR ($\geq$ 8 dB). This enables us to model the data as jittered zero-crossings, which in turn allows us to employ our modified Euclidean algorithms (MEAs) [23] and their least squares refinements [43] to the data. We estimate the (half-) period of a noisy sinusoid directly using the MEAs, which are methods for finding the greatest common divisor (gcd) of a noisy contaminated data set [23, 43]. This approach is motivated by the fact that, in the noise-free case, the gcd of a sparse set of the first differences of the zero crossing times is very highly likely to be the half-period of the sinusoid. We show that the method achieves a Cramer-Rao bound in a linear regression framework that arises naturally because of the high SNR assumption. Because the MEAs can successfully work with sparse data with jitter noise, this adds a new approach to solving zero-crossing problems that work when other methods break down. Moreover, given sufficient data, the proposed approach can tolerate outliers. We also note that our method can work even when the Tretter approximation does not hold. Our MEA with averaging identifies data in a tight cluster, and averages over these values. This has the effect of averaging the noise around the true value. We then employ the standard to this "averaged" data.

The journal paper includes two (analytic and geometric) derivations of the Tretter approximation, i.e., if $\mathrm{SNR}_s \gg 1$ (high SNR), then

$$s(n) = A \exp i(\omega_0 nT + \phi) + z(n) \approx A \exp(i(\omega_0 nT + \phi + \beta(n)))\,.$$

We also develop bounds to show why this approximation works for $\mathrm{SNR}_s \geq 8$ dB .

We have developed procedures for simple spectral analysis that combine filtering and number-theoretic based methods. The basic procedure is to first filter noisy time-series data, and extract zero-crossing data from the series. This data is then analyzed by using modifications of the Euclidean algorithm to isolate the fundamental period of the data, taking advantage of the fact that these algorithms are computationally straightforward, converge quickly, and are robust in that they are stable despite the presence of noise and arbitrary outliers. This gives a procedure for performing a basic spectrum analysis on extremely noisy time-series data.

We have also concentrated on the analysis of multiply periodic point processes [21, 24]. For these processes, we assume our data model is the union of M copies of (1), each with different periods, $k_{m,j}$'s and phases. We also assume that the $\eta_{m,j}$'s are zero-mean iid. We think of the data as being arrival times, and denote it by $\{\delta_{m,j}\}$. Assuming only minimal knowledge of the range of $\{\tau_m\}$, namely bounds T_L, T_U such that $0 < T_L \le \tau_m \le T_U$, we phase wrap the data by the mapping

$$\Phi_\rho(\delta_{m,j}) = \left\langle \frac{\delta_{m,j}}{\rho} \right\rangle = \frac{\delta_{m,j}}{\rho} - \left\lfloor \frac{\delta_{m,j}}{\rho} \right\rfloor, \tag{14}$$

where $\rho \in [T_L, T_U]$, and $\lfloor \cdot \rfloor$ is the floor function. Thus $\langle \cdot \rangle$ is the fractional part, and so $\left\langle \frac{t_j}{\rho} \right\rangle \in [0,1)$. We prove for almost every choice of ρ (in the sense of Lebesgue measure) that $\Phi_\rho(\delta_{m,j})$ is *essentially uniformly distributed* in the sense of Weyl. Moreover, the set of ρ's for which this is not true are fractions of $\{\tau_m\}$.

We then map the phase wrapped data by non-linear variations on the periodogram, e.g.,

$$F(\delta_{m,j}, \rho) = \sum_{m,j} \cos^{2l-1}(2\pi \frac{\delta_{m,j}}{\rho}) + i \sum_{m,j} \sin^{2l-1}(2\pi \frac{\delta_{m,j}}{\rho}), \tag{15}$$

for $l = 2, 3, \ldots$. We find

$$\max_\rho (\Re F - |\Im F|) \, .$$

This isolates the most prolific of the τ_m's. We prove this result using our variation of Weyl's Theorem. We can then subtract out this data, and repeat.

Our techniques have been applied to other data sets containing sparse noisy information generated by a periodic process, e.g., the geometric pattern generated by minefield placement [36], regularities in scheduled events, etc. The application to general time series data is straightforward. To detect regularities in a sparse, noisy lattice pattern (such as one gets from minefield data), we proceeded as follows. We first apply the Hough transform, which gives us one-dimensional data to analyze. This data is then processed by the modified Euclidean algorithm. The approach takes advantage of prior information on minefield spacing to eliminate points which are either spaced too closely or too far apart. This bounds are used to set thresholds in the MEA.

This approach can be refined. The first improvement would be to employ the least squares refinements [43] to the MEA. Also, the low complexity of the algorithm would allow it to be a component of an iterative scheme, in which the estimate of d is refined relative the data.

3 Multichannel Deconvolution

Deconvolution is one of the most general inverse problems. Results in this area are extremely useful, in that they have immediate application to not only theoretical but also applied mathematics. The theory of deconvolution presented in this report is contained in a larger group of results in the theory of residues of analytic functions and their generalities, for example, intersection varieties. These results have appeared in a series of papers by Berenstein, Gay, Yger et al. (see [2] – [19]), and can be interpreted as results in division problems, interpolation of analytic functions, analytic continuation, digital to analog conversion, and complexity theory. For deconvolution and other applications to signal and image processing, the work focuses on solving the general analytic Bezout equation, i.e., for given holomorphic f_i and ψ, solving for holomorphic g_i such that

$$\psi = f_1 \cdot g_1 + \ldots + f_n \cdot g_n. \tag{16}$$

In many situations, we want $\psi = \psi_\lambda$, with $\psi_\lambda \to 1$ as $\lambda \to \infty$ (ψ_λ is the transform of an approximate identity). Solutions to Bezout equations have yielded results in deconvolution, complexity theory, solutions to systems of PDE's, theorems about interpolation and continuation of analytic functions, and results in number theory (see [2] – [19]). We look to continue development of the general theory. However, in the context of this report, we have focused on the problem of deconvolution and application of the theory to signal and image processing. Our philosophy is that these specific results will still contribute to the development of the general theory, as the flow of ideas and results goes freely from theory to applications and back. Many of the PI's results on the subject appear in [11] – "Systems of convolution equations, deconvolution, Shannon sampling, and the wavelet and Gabor transforms," *SIAM Review* **36** (4), pp. 537–577 (1994) – [12] – "Exact multichannel deconvolution on radial domains," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, Vol. 3, 1865–1868 (1997). – [13] – "Sampling techniques for multichannel deconvolution," *1997 International Workshop on Sampling Theory and Applications*, 279–284 (1997). – [14] – "Modulation and sampling techniques for multichannel deconvolution," *Journal of Inverse and Ill-Posed Problems*, **8** (1), 1–41 (1998). Also see [15, 16, 17, 18, 19].

Linear, translation-invariant systems are modeled as follows. Assuming system impulse response function μ, they are represented by the convolution equation $s = f * \mu$. The output s may be an inadequate approximation of f, which motivates solving the convolution equation for f, i.e. deconvolving f from μ. If the function μ is time-limited (compactly supported) and non-singular, this deconvolution problem is ill-posed. A theory of solving such equations has been developed. It circumvents ill-posedness by using a *multichannel system.* If we overdetermine the signal f by using a system of convolution equations, $s_i = f * \mu_i$, $i = 1, \ldots, n$, the problem of solving for f is well-posed if the set of convolvers $\{\mu_i\}$ satisfies the *strongly coprime* condition. In this case, there exist compactly supported distributions (deconvolvers) ν_i , $i = 1, \ldots, n$ which solve $1 = \widehat{\mu}_1 \cdot \widehat{\nu}_1 + \ldots + \widehat{\mu}_n \cdot \widehat{\nu}_n$, the Bezout equation. Transforming, we get $\delta = \mu_1 * \nu_1 + \ldots + \mu_n * \nu_n$, which in turn gives $f = s_1 * \nu_1 + \ldots + s_n * \nu_n$. The development of strongly coprime systems rests on a special non-overlapping alignment of the zeros of $\{\widehat{\mu}_i\}$, given as a number theoretic condition. The mutichannel systems are analogous to systems of relatively prime congruences.

We have continued the development of the theory. Particular items include:

i.) Extended the theory to convolvers which can be modeled by B-splines [18] and those systems (e.g., optical) which work in radial domains [12]. In particular, we now can apply the theory in general imaging systems (e.g., for sub-pixel resolution), medical imaging (e.g., strongly coprime x-ray systems), radar (e.g., strongly coprime chirp pulses), etc. Using duality, we can also use the theory to create filters.

ii.) Developed a basis in which to develop the multichannel theory which is more amenable to discretization [18].

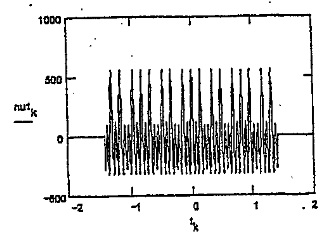
iii.) Coordinated the theory with filtering systems for the removal of noise and/or other unwanted information. This was achieved by combining multichanel deconvolution with Wiener filtering [13, 19].

iv.) Used the theory as a model for developing new sampling schemes. Solutions to the analytic Bezout equation associated with certain multichannel deconvolution problems are interpolation problems on unions of non-commensurate lattices. These solutions provide insight into how we developed general sampling schemes on properly chosen non-commensurate lattices. We have developed specific examples of non-commensurate sampling lattices, and used a generalization of B. Ya. Levin's sine-type functions to develop interpolating formulae on these lattices [17]. We are exploring an extension of these ideas to create non-commensurate systems of wavelet bases.

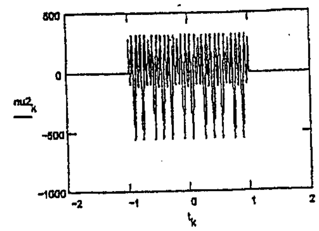
The work on multichannel deconvolution is certainly a new perspective on this class of inverse problems. Items **(i.)** – **(iii.)** provide a theoretical basis for multichannel systems in signal and image processing. This work is indirectly applicable to Navy technologies in that it provides an outline for the development of active and passive remote sensing systems which gather all possible information, new techniques for making filters, etc. The multichannel approach turns ill-posed problems into well-posed ones, and provides a framework for solution by linear methods. Moreover, the success of our simulation work gives proof that it would not be difficult to develop actual multichannel systems. The sampling schemes [17] also provide a theoretical basis for extending bandwith. This may be useful in, for example, A to D conversion systems.

4 A Direct Yet Fundamental Result on Derivatives

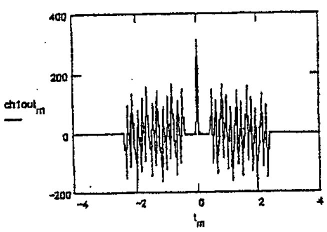
The paper *When Does a Positive Derivative Guarantee Monotonicity? (Some New Thoughts on the Classical Theory)*, [20], discusses one of the most fundamental and useful results from calculus. It is well known that if a function f is differentiable with positive derivative Df on an interval, then it is increasing there. If we



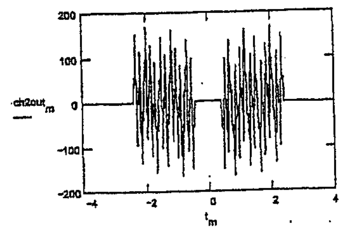
2a. Channel 1 Deconvolver (v_1).



2a. Channel 2 Deconvolver (v_2).

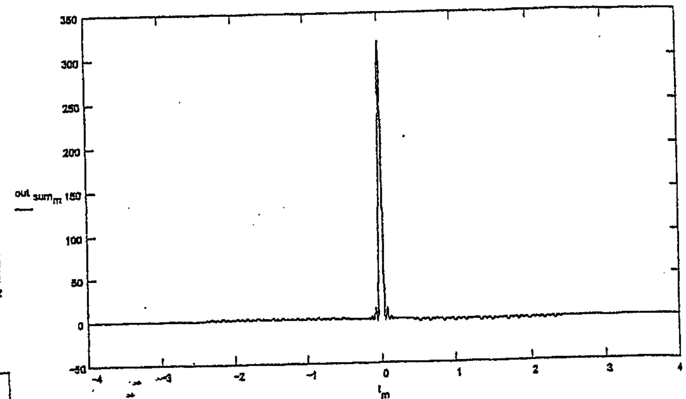


2c. Channel 1 output ($\Psi * \mu_1 * v_1$)

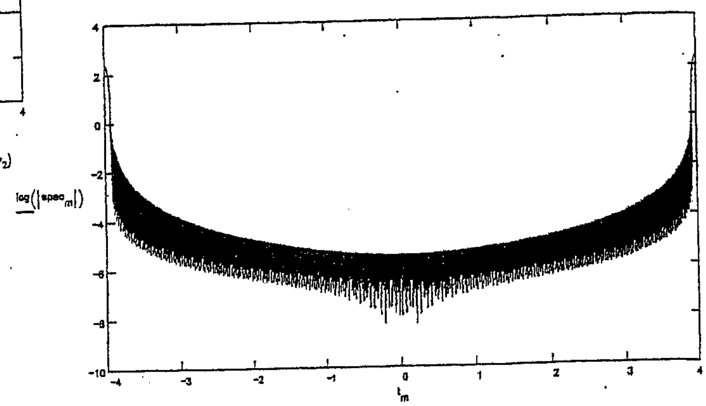


2d. Channel 2 output ($\Psi * \mu_2 * v_2$)

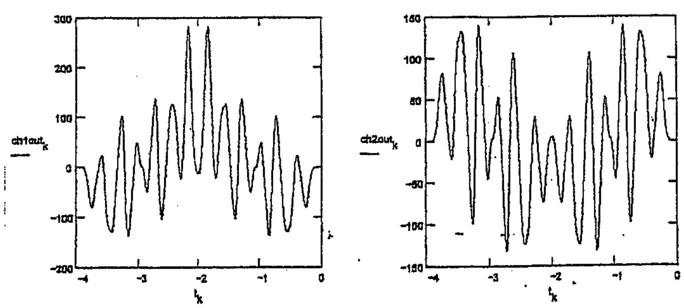
Figure 2. C^∞ Function (Ψ) Approximate Deconvolver System.



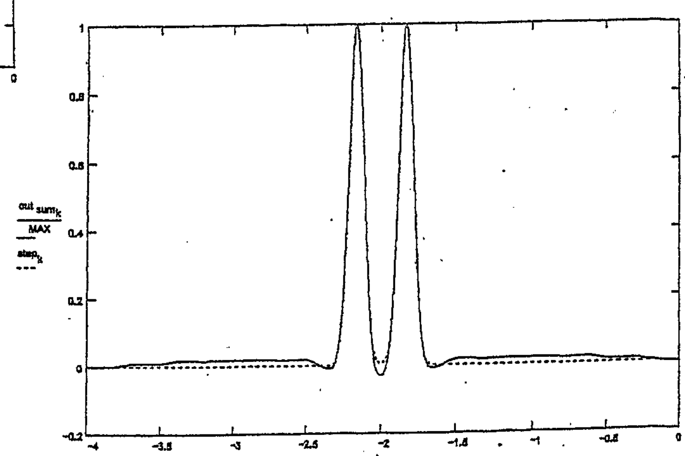
2e. $\tilde{\Psi} = \Psi * \mu_1 * v_1 + \Psi * \mu_2 * v_2$, the resulting approximate delta functional (time domain).



2f. $\tilde{\Psi} = \Psi * \mu_1 * v_1 + \Psi * \mu_2 * v_2$, the resulting approximate delta functional (frequency domain).



3a. The response of each channel, in the two channel system of Fig. 1, to two Gaussian pulses (denoted as the input signal, f).



3b. System transfer response, for the two channel system of Fig. 1, when the input function is the sum of two time shifted Gaussian pulses. The dotted line is the input function, the solid line is the deconvolved function.

Figure 3. Transfer Response for the System.

assume f is continuous on $[a,b]$ and differentiable on (a,b), then the result follows directly from the Mean Value Theorem. The paper asks about generalizations of this result. In particular, we obtain, given certain assumptions, the most straightforward yet most general theorem possible. We assume throughout the paper that all of the functions considered in connection to the investigation are continuous. If we do not assume continuity, we can vary behavior at a single point, thus easily changing monotonicity. The work boiled down into answering two questions.

If we weaken the hypotheses of the Mean Value Theorem,
does $Df > 0$ imply that f is increasing ?

As one would expect, the answer to our first question is no. To see this, we assumed, in addition to continuity, only that $Df > 0$ almost everywhere (*a.e.*), i.e., except on a set of Lebesgue measure zero. This condition was too weak, for we construct counterexamples to show that the assumption $Df > 0$ *a.e.* does not imply that f is increasing.

We develop our counterexamples using the *Cantor-Lebesgue function* $\mathcal{C}$, which is defined as follows. Let C denote the Cantor middle thirds set, and for $x \in [0,1]$, consider its ternary expansion $x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$. Let

$$n = n(x) = \begin{cases} \min\{k : a_k = 1\} & \text{if } x \notin C \\ \infty & \text{if } x \in C\,. \end{cases}$$

Then

$$\mathcal{C}(x) = \sum_{k=1}^{n-1} \frac{a_k}{2^{k+1}} + \frac{1}{2^n}\,.$$

$\mathcal{C}$ is a continuous not strictly monotone increasing function on [0,1] with $\mathcal{C}(0) = 0$ and $\mathcal{C}(1) = 1$ that has zero derivative except on the Cantor set, i.e., $D\mathcal{C} = 0$ *a.e.*

We can now construct our example. Let

$$f(x) = \frac{x}{2} + \mathcal{C}(1-x)\,.$$

Then f is a continuous mapping on $[0,1]$ with $Df = \frac{1}{2}$ *a.e.* However, $f(0) = 1 > \frac{1}{2} = f(1)$. The function f is not increasing. Here, the "risers" of the Cantor-Lebesgue function, which occur over the Cantor middle-thirds set, allow the function f to "flow" against the derivative. Measure zero gives too much room, allowing for this flow. We can, of course, produce other such functions, such as, for fixed $\alpha \in (0,1)$, $f_\alpha(x) = \alpha x + \mathcal{C}(1-x)$ and $g_\alpha(x) = \alpha x - \mathcal{C}(x)$.

This naturally led us to our second question.

Assuming that f is continuous on $[a,b]$ and differentiable on (a,b) *a.e.*,
what is the weakest additional condition we can impose on Df
which guarantees that f is monotone increasing ?

If we assume that the set on which the hypotheses fail is countable, we prove the following. We refer to this as the *Countable Exceptional Set Theorem.*

Theorem 4.1 *Let f be a continuous function on $[a,b]$ and suppose that $Df > 0$ on $[a,b] \setminus S$, where S is countable. Then $f(a) < f(b)$.*

A reading of Saks' classic treatise *Theory of the Integral* gives that this result is generalized by the Goldowsky-Tonelli Theorem. (Also see Kannan and Krueger's *Advanced Analysis on the Real Line.*)

Theorem 4.2 (Goldowsky-Tonelli) *Let f be a continuous function on $[a,b]$ and suppose that Df exists (finite or infinite) on $[a,b] \setminus S$, where S is countable. Also suppose that $Df \geq 0$ a.e. on $[a,b]$. Then f is a non-decreasing function on $[a,b]$.*

We give a new proof of Goldowsky-Tonelli in our paper. This still, however, does not answer our second question.

We give a surprising answer in [20]. We show that, in a very natural sense, Theorem 4 is the best possible result. Goldowsky-Tonelli is a vacuous extension of the Countable Exceptional Set Theorem, in that there are no additional functions to which Goldowsky-Tonelli applies but Countable Exception does not apply. Goldowsky-Tonelli, however, requires less information. Given a function satisfying the hypotheses of Goldowsky-Tonelli, one needs Goldowsky-Tonelli to show that it satisfies the hypotheses of Countable Exception. We develop our proof by proving the next two theorems, and using a classical result of Alexandroff and Hausdorff.

Theorem 4.3 *If g is continuous and the set of points where Dg is not positive (or fails to exist) does not contain a perfect set, then g is non-decreasing. However, for any set that does contain a perfect set, there are counterexamples.*

Theorem 4.4 *The set of points where a continuous function fails to have a positive derivative is a Borel set.*

Theorem 4.5 (Alexandroff and Hausdorff) *Any uncountable Borel set contains a perfect set.*

Putting these theorems together, we show that, in fact, Theorem 4 is best possible. There is no condition between countable and measure zero that guarantees monotonicity.

5 Appendix: Publications and Awards

5.1 Publications

The following gives a listing of papers whose research is partially supported by the grant.

1. Casey, S. D., Berenstein C. A., and Walnut, D. F., "Exact multichannel deconvolution on radial domains," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, Vol. 3, 1865–1868 (1997).

2. Casey, S. D., "Sampling techniques for multichannel deconvolution," *1997 International Workshop on Sampling Theory and Applications*, 279–284 (1997).

3. Casey, S. D., "Modulation and sampling techniques for multichannel deconvolution," *Journal of Inverse and Ill-Posed Problems*, **8** (1), 1–41 (1998).

4. Casey, S. D., "Sampling issues in least squares, Fourier analytic, and number theoretic methods in parameter estimation," *31st Annual Asilomar Conference on Signals, Systems, and Computers*, 453–457 (1998).

5. Casey, S. D., "New classes of Berenstein deconvolvers and their applications to filter design," 46 pp., preprint.

6. Casey, S. D., "Optimal multichannel deconvolution in a noisy environment: deconvolving Wiener filters," 28 pp., preprint.

7. Casey, S. D., and Holzsager, R. A., "When does a positive derivative guarantee monotonicity?" submitted to *The American Mathematical Monthly*, 17 pp. (1998).

8. Casey, S. D., Berenstein C. A., and Walnut, D. F., "Systems of convolution equations, deconvolution, Shannon sampling, and the wavelet and Gabor transforms," invited research monograph for *SIAM Monographs on Mathematical Modeling and Computation* (in progress).

9. Casey, S. D., Berenstein C. A., and Walnut, D. F., "Exact multichannel deconvolution," invited survey article for *Advances in Imaging and Electron Physics* (in progress).

10. Casey, S. D., and Walnut, D. F., "Sampling on unions of non-commensurate lattices via real and complex interpolation theory," invited article for research monograph on sampling theory in *Birkhauser Research Monographs* (in progress).

11. Lake, D., Sadler, B., and Casey, S., "Detecting regularity in minefields using collinearity and a modified Euclidean algorithm, *Proc. SPIE*, **3079**, 8 pp. (1997).

12. Sadler, B. M., and Casey, S. D., "On periodic pulse interval analysis with outliers and missing observations," to appear in *IEEE Transactions on Signal Processing*, 31 pp. (1998).

13. Sadler, B. M., and Casey, S. D., "Frequency estimation via sparse zero crossings," submitted to *IEEE Transactions on Signal Processing*, 11 pp. (1998).

5.2 Awards

The PI received the following award during the grant period.

1. 1998 American University Faculty Award for Outstanding Teaching.

References

[1] Bartlett, M. S., "The spectral analysis of point processes," *Journal of the Royal Statistical Society B*, **25** (2), 264–280 (1963).

[2] Berenstein, C. A., and Yger, A., "Effective Bezout identities in $\mathbf{Q}[z_1, \ldots, z_n]$," *Acta Mathematica* **166**, 69–120 (1991).

[3] Berenstein, C. A. and Yger, A., "Analytic Bezout identities," *Advances in Mathematics* **10**, 51–74 (1989).

[4] Berenstein, C. A., Gay, R., and Yger A., "Inversion of the local Pompeiu transform," *J. Analyze Math.* **54**, 259–70 (1990).

[5] Berenstein, C. A., and Yger, A., "Le probleme de la deconvolution," *J. Funct. Anal.* **54**, 113–60 (1983).

[6] Berenstein C. A., Yger, A., and Taylor, B. A., "Sur quelques formules explicites de deconvolution," *Journal of Optics (Paris)* **14**, 75–82 (1983).

[7] Berenstein C. A., and Patrick, E. V., "Exact deconvolution for multiple convolution sensors–an overview plus performance characterizations for imaging sensors," *Proceedings of the IEEE (Special Issue on Multidimensional Signal Processing)*, **78** (4), 723–34 (1990).

[8] Berenstein C. A., Patrick, E. V., and Casey, S. D., "Systems of convolution equations, deconvolution, and wavelet analysis," *Systems Research Center whitepaper*, 25 pp. (1991).

[9] Bloomfield, P., "Spectral analysis with randomly missing observations," *Journal of the Royal Statistical Society B*, **32** (3), 369–380 (1970).

[10] Casey, S. D., "Deconvolution and Gabor and wavelet signal analysis," *Proceedings of the EFTF*, 262–274 (1991).

[11] Casey, S. D., and Walnut, D. F., "Systems of convolution equations, deconvolution, Shannon sampling, and the wavelet and Gabor transforms," *SIAM Review* **36** (4), 537–577 (1994).

[12] Casey, S. D., Berenstein C. A., and Walnut, D. F., "Exact multichannel deconvolution on radial domains," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, Vol. 3, 1865–1868 (1997).

[13] Casey, S. D., "Sampling techniques for multichannel deconvolution," *1997 International Workshop on Sampling Theory and Applications*, 279–284 (1997).

[14] Casey, S. D., "Modulation and sampling techniques for multichannel deconvolution," *Journal of Inverse and Ill-Posed Problems*, **8** (1), 1–41 (1998).

[15] Casey, S. D., Berenstein C. A., and Walnut, D. F., "Systems of convolution equations, deconvolution, Shannon sampling, and the wavelet and Gabor transforms," invited research monograph for *SIAM Monographs on Mathematical Modeling and Computation* (in progress).

[16] Casey, S. D., Berenstein C. A., and Walnut, D. F., "Exact multichannel deconvolution," invited survey article for *Advances in Imaging and Electron Physics* (in progress).

[17] Casey, S. D., and Walnut, D. F., "Sampling on unions of non-commensurate lattices via real and complex interpolation theory," invited article for research monograph on sampling theory in *Birkhauser Research Monographs* (in progress).

[18] Casey, S. D., "New classes of Berenstein deconvolvers and their applications to filter design," 46 pp., preprint.

[19] Casey, S. D., "Optimal multichannel deconvolution in a noisy environment: deconvolving Wiener filters," 28 pp., preprint.

[20] Casey, S. D., and Holzsager, R. A., "When does a positive derivative guarantee monotonicity?" submitted to *The American Mathematical Monthly*, 17 pp. (1998).

[21] Casey, S. D., "Sampling issues in least squares, Fourier analytic, and number theoretic methods in parameter estimation," *31st Annual Asilomar Conference on Signals, Systems, and Computers*, 453–457 (1998).

[22] Casey, S. D., and Sadler, B. M., "A modified Euclidean algorithm for isolating periodicities from noisy data," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '95)*, Vol. 3, 1764–1767 (1995).

[23] Casey, S. D., and Sadler, B. M., "Modifications of the Euclidean algorithm for isolating periodicities from a sparse set of noisy measurements," *IEEE Transactions on Signal Processing*, 44 (8), 2260–2272 (1996).

[24] Casey, S. D., and Sadler, B. M., "Number theoretic methods in parameter estimation," *Proceedings of IEEE Workshop on Statistical Signal and Array Processing*, 406–409 (1996).

[25] Finette, S., and Mignerey, P., "Multichannel deconvolution of an acoustic transient in an oceanic waveguide," *Jour. Acoustical Soc.* **92** (1), 351–364 (1992).

[26] Hörmander, L., *The analysis of linear partial differential operators I (distribution theory and Fourier analysis), second edition*, Springer-Verlag, New York (1990).

[27] Hörmander, L., "Generators for some rings of analytic functions," *Bull. American Math. Soc.*, **73**, 943–949, (1967).

[28] Ireland, K., and Rosen, M., *A Classical Introduction to Modern Number Theory*, Springer-Verlag, New York (1982).

[29] Jain, A. K., *Fundamentals of Digital Image Processing*, Prentice-Hall, New Jersey (1989).

[30] Jones, R. H., "Spectral analysis with regularly missed observations," *Annals of Mathematical Statistics* **33** 455–461 (1962).

[31] Kay, S. M., *Fundamentals of Statistical Signal Processing*, Prentice-Hall, New Jersey (1993).

[32] Kay, S. M., and Sudhaker, R., "A zero crossing-based spectrum analyzer," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **34** (1), 96–104 (1986).

[33] Kedem, B., *Time Series Analysis by Higher Order Crossings*, IEEE Press, New York (1994).

[34] Knuth, D. E., *The Art of Computer Programming, Volume 2: Seminumerical Algorithms (Second Edition)*, Addison-Wesley, Reading, Massachusetts (1981).

[35] Knuth, D. E., *The Art of Computer Programming, Volume 3: Sorting and Searching*, Addison-Wesley, Reading, Massachusetts (1973).

[36] Lake, D., Sadler, B. M., and Casey, S. D., "Detecting regularity in minefields using collinearity and a modified Euclidean algorithm, *Proc. SPIE*, **3079**, 8 pp. (1997).

[37] Leveque, W. J., *Topics in Number Theory, Volumes 1 and 2*, Addison-Wesley, Reading, Massachusetts (1956).

[38] Papoulis, A., *Signal analysis*, McGraw-Hill, New York (1977).

[39] Papoulis, A., *Probability, Random Variables, and Stochastic Processes (Third Edition)*, McGraw-Hill, New York (1991).

[40] Parzen, E., "On spectral analysis with missing observations and amplitude modulation," *Sankhya* (Ser. A), **25**, 383–392 (1963).

[41] Sadler, B. M., and Casey, S. D., "PRI analysis from sparse data via a modified Euclidean algorithm," *29th Annual Asilomar Conference on Signals, Systems, and Computers*, pp. 161–168 (1995).

[42] Sadler, B. M., and Casey, S. D., "Frequency estimation via sparse zero crossings," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '96)*, Vol. 5, 2990–2993 (1996).

[43] Sadler, B. M., and Casey, S. D., "On periodic pulse interval analysis with outliers and missing observations," to appear in *IEEE Transactions on Signal Processing*, 31 pp. (1998).

[44] Sadler, B. M., and Casey, S. D., "Frequency estimation via sparse zero crossings," submitted to *IEEE Transactions on Signal Processing*, 11 pp. (1998).

[45] Schroeder, M. R., *Number Theory in Science and Communication (Second Edition)*, Springer-Verlag, Berlin (1986).

[46] Shannon, C. E., "Communication in the presence of noise," *Proc. IRE* **37**, 10–21 (1949).

[47] Wiener, N. *The Fourier integral and certain of its applications*, MIT Press, Cambridge, MA (1933).

[48] Wiener, N. *Extrapolation, interpolation, and smoothing of stationary time series*, MIT Technology Press, Cambridge, MA (1949).

[49] Wiley, R. G., *Electronic Intelligence: The Analysis of Radar Signals (Second Edition)*, Artech House, Norwood, Massachusetts (1993).